



The Frozen Face Effect: Why Static Photographs May Not Do You Justice

Citation

Post, Robert B., Jason Haberman, Lica Iwaki, and David Whitney. 2012. The frozen face effect: Why static photographs may not do you justice. *Frontiers in Psychology* 3:22.

Published Version

doi://10.3389/fpsyg.2012.00022

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:8457941>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)



The frozen face effect: why static photographs may not do you justice

Robert B. Post¹, Jason Haberman^{1,2,3}, Lica Iwaki^{1,2} and David Whitney^{2,4*}

¹ Department of Psychology, University of California at Davis, Davis, CA, USA

² Center for Mind and Brain, University of California at Davis, Davis, CA, USA

³ Department of Psychology, Harvard University, Cambridge, MA, USA

⁴ Department of Psychology, University of California at Berkeley, Berkeley, CA, USA

Edited by:

J. Toby Mordkoff, University of Iowa, USA

Reviewed by:

Erin Heerey, Bangor University, UK

Marc Grosjean, Leibniz Research Centre for Working Environment and Human Factors, Germany

*Correspondence:

David Whitney, Department of Psychology, University of California at Berkeley, 3210 Tolman Hall, Berkeley, CA 94720, USA.
e-mail: dwhitney@berkeley.edu

When a video of someone speaking is paused, the stationary image of the speaker typically appears less flattering than the video, which contained motion. We call this the frozen face effect (FFE). Here we report six experiments intended to quantify this effect and determine its cause. In Experiment 1, video clips of people speaking in naturalistic settings as well as all of the static frames that composed each video were presented, and subjects rated how flattering each stimulus was. The videos were rated to be significantly more flattering than the static images, confirming the FFE. In Experiment 2, videos and static images were inverted, and the videos were again rated as more flattering than the static images. In Experiment 3, a discrimination task measured recognition of the static images that composed each video. Recognition did not correlate with flattery ratings, suggesting that the FFE is not due to better memory for particularly distinct images. In Experiment 4, flattery ratings for groups of static images were compared with those for videos and static images. Ratings for the video stimuli were higher than those for either the group or individual static stimuli, suggesting that the amount of information available is not what produces the FFE. In Experiment 5, videos were presented under four conditions: forward motion, inverted forward motion, reversed motion, and scrambled frame sequence. Flattery ratings for the scrambled videos were significantly lower than those for the other three conditions. In Experiment 6, as in Experiment 2, inverted videos and static images were compared with upright ones, and the response measure was changed to perceived attractiveness. Videos were rated as more attractive than the static images for both upright and inverted stimuli. Overall, the results suggest that the FFE requires continuous, natural motion of faces, is not sensitive to inversion, and is not due to a memory effect.

Keywords: face perception, static images, dynamic images, attractiveness, fluency

INTRODUCTION

We have often observed that when a video of someone speaking is paused, the stationary image of the speaker typically appears less flattering than the preceding video. We refer to this phenomenon as the frozen face effect (FFE).

The FFE may be related to observations in prior research that facial motion may influence ratings of attractiveness. For example, using computer-generated animations of faces, Knappmeyer et al. (2002) report that the addition of motion to computerized faces positively influenced attractiveness judgments. Morrison et al. (2007) report that animating androgynous line drawings of faces can influence attractiveness ratings. Specifically, attractiveness was found to correlate with the total amount of movement in female, but not male faces. This study did not compare the animations to static faces, however. In contrast to the studies using computer-generated or cartoon-like faces, Rubenstein (2005) compared attractiveness of video clips of a person speaking and a stationary image selected from the video clip (on the basis of neutral appearance). In this study, there was no difference in attractiveness ratings between the stationary and dynamic facial stimuli.

In this paper we sought to empirically demonstrate the FFE, and examine its possible mechanisms. In contrast to past work, we presented more ecologically valid stimuli, including videos of individuals speaking in naturalistic settings such as news programs, talk shows, and interviews. Speakers included individuals who were either famous or not, and speaking a variety of languages.

EXPERIMENT 1

Experiment 1 was designed to quantify the FFE and measure differences in how subjects rated flattery of the video and static frame stimuli.

METHOD

Ethics statement – Written informed consent was obtained for all participants, and UC Davis' Institutional Review Board granted approval for all research.

Subjects

Seven undergraduate students (two males, five females), aged 21–23 participated. All but one observer were naïve to the hypotheses of the research.



FIGURE 1 | Examples of static images derived from video clips.

Stimuli

The stimuli were 40, 2 s video clips (20 unique individuals, 4 s of video for each individual divided into two equal 2 s clips), and all static frames contained within each video. Video clips of individuals speaking in naturalistic settings were sampled from the internet, divided into 2 s clips, and saved in Quicktime format at 15 frames per second (fps). Although familiarity (i.e., with celebrities) was not controlled, anecdotally participants reported recognizing less than one-third of the individuals. Static stimuli were created by extracting all frames from each 2 s clip, which were standardized at 528×431 pixels. Video clips were then recreated from the static frames at 15 fps and muted to eliminate auditory cues. Thus a total of 40 2 s videos and 1200 static frames were created and presented as stimuli. Examples of static frames used in the study are shown in **Figure 1**.

Procedure

Subjects were seated in a dark sound-dampened room. Stimuli (all $11.6^\circ \times 9.50^\circ$) were presented on a Sony CRT (Sony Multi-scan G520, 21", 1600 \times 1200, 85 Hz refresh) at a viewing distance of 65 cm. All videos and static frames were presented in random order within the same session. Video stimuli were presented for the 2-s duration of each video, after which observers saw only a fixation cross until a response was made. Static stimuli were also presented until a subject response was received. A random dot mask was presented between each trial. After each image was presented, subjects gave flattery ratings for the person in each stimulus using a seven-point scale (1 = least flattering, 7 = most flattering). The next stimulus was presented immediately after the flattery rating was made. Attractiveness ratings have been widely used in previous research (e.g., Knappmeyer et al., 2002; Rubenstein, 2005; Morrison et al., 2007), however, attractiveness is not the optimal dependent measure for our task. While attractiveness is a perceptual property invariant to context, we were interested in relative differences as a function of stimulus type – hence our use of flattery as our dependent measure. To illustrate, one can imagine a scenario in which person A is known and believed to be

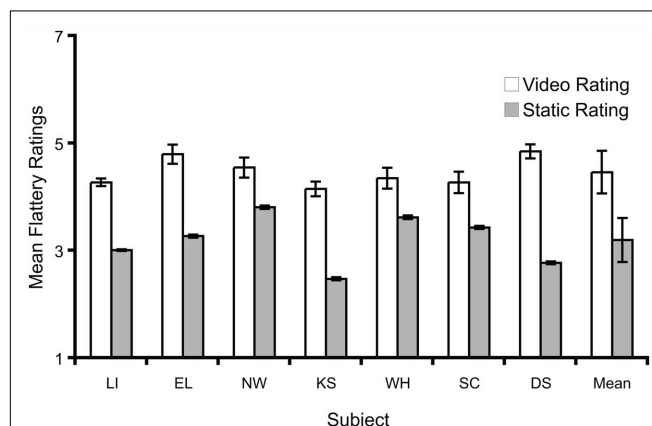


FIGURE 2 | Flattery ratings in Experiment 1 for the person in each video or static frame using a seven-point scale (1 = least flattering, 7 = most flattering). Error bars denote \pm SEM.

attractive, but appears in a photograph that is not particularly flattering. By utilizing flattery ratings, we offset differences in absolute attractiveness across our stimuli and reduced subject confusion about what counts as “attractive.” Nevertheless, we conducted a control experiment (Experiment 6) that used attractiveness ratings, to confirm that the same results hold with both flattery and attractiveness ratings.

RESULTS AND DISCUSSION

Overall, flattery ratings of the videos were significantly higher than average flattery ratings of static images derived from the videos [$t(6) = 6.40, p < 0.001, \eta = 0.87$]. As seen in **Figure 2**, this pattern was obtained for each of the seven subjects. We assessed inter-rater reliability of the static image ratings by using Kendall's W . For each observer, we averaged the ratings of the 30 images composing a given movie. Thus, we compared rating consistency for 40 stimuli across the seven subjects. The inter-rater reliability was significant [$W = 0.638, \chi^2(6) = 153, p < 0.001$], justifying the use

of a *t*-test and other comparable group assessments. It additionally demonstrates that subjects were not lapsing or repeatedly making the same response as they rated the static images.

The results demonstrate a very strong effect for each of the seven subjects. On average, subjects rated 91% of the static frames as less flattering than the movies they composed (excluding images rated the same as videos; across subjects this translated to 6254 out of 6844 static images being rated as less flattering than the corresponding video). This pattern held for each subject; the least significant subject still rated 73% of the static images as less flattering than the corresponding movies (652 out of 882), a highly significant effect [$\chi^2(1) = 201, p < 0.0001$]. The consistency of the FFE for the vast majority of the static images and for each of the 40 movie stimuli shows that familiarity with certain faces (which occurred on less than one-third of the stimuli) was not necessary for the FFE.

This finding is consistent with the results of Knappmeyer et al. (2002), but contrary to the results of Rubenstein (2005) who found no influence of facial motion on ratings of attractiveness.

A regression analysis of the mean flattery ratings for the static images against flattery ratings for the corresponding video clips indicated a statistically significant relationship [average $r = 0.58$ (converted from averaged fisher *z*-scores), least significant subject: $r = 0.35, p = 0.025$]. This contrasts with Rubenstein (2005), who reported no correlation between attractiveness ratings of static and dynamic formats of the same face ($r = 0.19, p = 0.26$).

A number of possibilities exist for this discrepancy, including differences in the stimuli. Rubenstein (2005) was careful to use neutral static images, in which the depicted individual was staring straight ahead. Here, we used natural movies of speaking individuals where the expression, gaze, and head orientation could vary.

The results of Experiment 1 provide strong empirical support for the existence of the FFE. The rest of this paper explores the properties, limitations, and possible mechanisms of the FFE.

EXPERIMENT 2: THE INVERTED FFE

Prior research has shown that face recognition is strongly disrupted by facial inversion (e.g., Yin, 1969; Farah et al., 1997). Additionally, artificial manipulations of faces that are easily detected when they are presented upright, become much less perceptible when they are presented inverted (Thompson, 1980). These studies might predict that the FFE is either strongly decreased or eliminated with facial inversion. Experiment 2 was therefore designed to test this hypothesis.

METHOD

Subjects

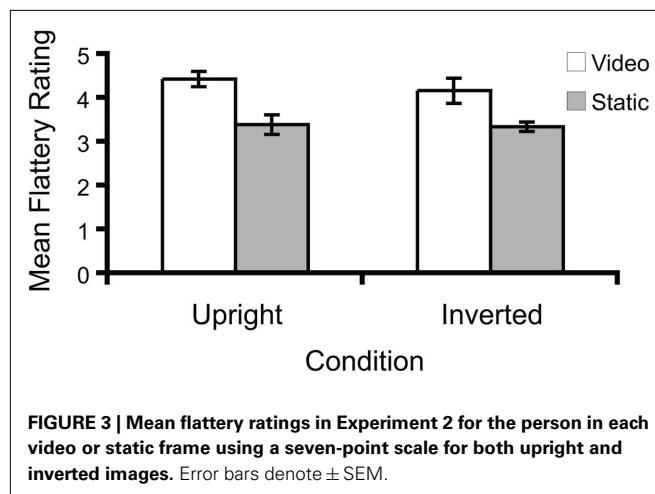
Four of the subjects from Experiment 1 participated in the study.

Stimuli

The stimuli were the 40 videos and 1200 static frames created for Experiment 1, and inverted copies of them.

Procedure

The procedure was similar to that of Experiment 1, with all videos and static frames presented in random order and subjects



providing flattery ratings for the person in each video or static frame using a seven-point scale.

RESULTS AND DISCUSSION

Mean flattery ratings for both upright and inverted static and video images are shown in **Figure 3**. It is apparent in the figure that the FFE persisted despite facial inversion. A two (image: static vs. video) by two (orientation: upright vs. inverted) repeated-measures ANOVA revealed significant main effect of image ($F_{1,3} = 194.90, p < 0.005$), and a marginally significant effect of orientation ($F_{1,3} = 10.00, p = 0.05$). Thus, the mean FFE – the difference scores (1.04 for upright videos minus upright static images; 0.82 for inverted videos minus inverted static images) were significantly higher than 0. Interestingly, the interaction between inversion and stimulus type was not significant ($F_{1,3} = 0.19, p > 0.05$), suggesting that the magnitude of the FFE was not differentially affected by changing the orientation of the faces.

The results obtained with the upright stimuli replicate those of Experiment 1, further supporting the existence of the FFE. Videos were again rated as more flattering than the mean of the static images composing them. The finding that the FFE remains strong with facial inversion (about 80% of that obtained with upright stimuli) is unexpected, given that inversion is highly disruptive to facial recognition (e.g., Yin, 1969). It is also surprising in the context of studies demonstrating that artificial manipulations of faces that are easily detected when they are presented upright, become much less perceptible when they are presented inverted (Thompson, 1980). The results suggest that the FFE may not hinge upon configural or holistic information about faces.

EXPERIMENT 3: BETTER MEMORY FOR FLATTERING FACES?

One possible explanation for the FFE is that memory is better for distinct face images than for unremarkable face images. That is, although the video clips contain images that differ widely in how flattering they are, the more extreme ones (either particularly flattering or unflattering) may be remembered more, which could, in theory, bias the mean of the remembered faces toward one end of the flattery scale. Experiment 3 was designed to test this possibility.

METHOD

Subjects

Five of the subjects from Experiment 1 participated in the study.

Stimuli

The stimuli were the same 40 videos and 1200 static frames from Experiment 1.

Procedure

An ABX discrimination task was used to measure recognition of the static faces that composed each video. On each trial, a video was presented, followed immediately by two static images: a target image, and a lure image. The target image was a member of the 2-s video presented previously. As described in Experiment 1, two video clips were created for each individual (for each individual, a 4-s clip was cut in half, producing two temporally consecutive 2-s clips). The lure image was a randomly selected frame from the other 2-s video of the same individual, which was either immediately preceding or following the target video. Subjects were required to judge which of the two static images was a member of the previously seen video. Each frame of each video was a target once, for a total of 1200 trials.

RESULTS AND DISCUSSION

For each static image, a recall accuracy statistic was generated for each trial wherein if the target image was selected, the value 1 was assigned and if the lure image was selected the value 0 was assigned. The average accuracy for each set of 30 static images that composed each video was calculated, for a total of 40 means per subject. The aggregate mean performance across subjects and across videos was 0.71. However, the critical test here is whether there exists a relationship between accuracy and flattery ratings at the individual subject level. To examine this, we computed Spearman's Correlation Coefficient (non-parametric test). **Figure 4A** depicts the relationship between mean image flattery ratings and recall accuracy for one representative observer. No subject's correlation was significantly different from 0 (bootstrap test, $p > 0.05$ for all subjects; **Figure 4B**), suggesting no relationship between image flattery rating and probability of image recognition.

The results demonstrate that subjects did not simply remember faces that were more distinct (either more or less flattering). Accordingly, the FFE cannot be explained by better memory for flattering faces.

EXPERIMENT 4: ENSEMBLE RATINGS

Another possible cause of the FFE is that flattery ratings are influenced by the amount of information presented. Specifically, each video clip (composed of 30 static images) contains more information than is contained in any individual static image. Experiment 4 was designed to determine whether the FFE occurs because there is more information available in the video stimuli than in the static frame stimuli. To test this, we compared flattery ratings for video clips with ratings for an ensemble stimulus – an array of all the static images contained within the video clip displayed simultaneously.

METHOD

Subjects

Four of the subjects from Experiment 1 participated.

Stimuli

Static ensemble displays were created from all 30 static frames contained within each video. All frames were presented simultaneously within a six (horizontal) by five (vertical) grid, with the location of each frame randomly assigned.

Procedure

Ensemble displays were presented for 2 s to match the duration of the videos used in Experiment 1 and subjects gave flattery ratings in the same manner as the other studies.

RESULTS AND DISCUSSION

Mean flattery ratings were calculated for each of the ensemble stimuli. These are presented in **Figure 5**, together with the mean ratings for both the corresponding individual static stimuli and video clips from Experiment 1. It is apparent in the figure that the ratings for the ensemble stimuli were consistent with those given to the mean of the component individual images in Experiment 1. However, the ratings for the video

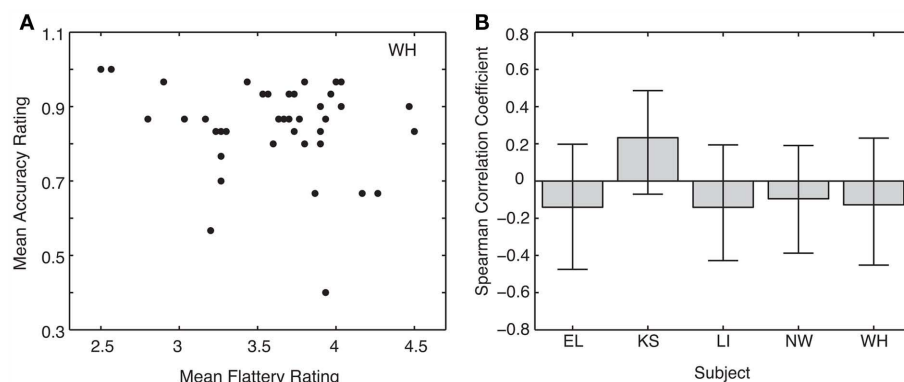


FIGURE 4 | (A) Mean recall accuracy as a function of mean flattery ratings in Experiment 3 for subject WH. Spearman correlations **(B)** were calculated

from these values for each individual subject. Error bars are 95% confidence intervals derived from 1000 bootstrapped estimates.

clips are clearly higher than either the ensemble stimuli ratings or the component image ratings obtained in Experiment 1. A one-way ANOVA revealed a marginally significant difference among the three conditions depicted in **Figure 5** ($F_{2,9} = 4.23$, $p = 0.051$).

An analysis of the individual subject data revealed that there was a significant correlation between static image ratings and ensemble ratings [average subject: $r = 0.88$ (converted from averaged Fisher z -scores); least significant subject: $r = 0.48$, $df = 39$, $p = 0.0015$], suggesting that individuals responded consistently between the two stimulus types. That is, a group of static images was rated similarly, whether that group was presented all at once as an ensemble or presented as a series of individual images. Importantly, it also shows that subjects were not lapsing or repeatedly making the same response during the experiment.

The finding that mean ensemble stimulus ratings were very similar to those for the component static images, and that the two measures correlated highly, are consistent with the results of Haberman and Whitney (2007, 2009), who reported that emotionality ratings for ensemble stimuli correlate highly with the mean of the ratings for the component images. The FFE is interesting in this context, as the video ratings are higher than those for either the individual static images or the ensemble stimuli. Therefore, with moving images, perceived flattery is not predicted entirely by extraction of the mean flattery ratings from the component static images.

The high correlation between individual static ratings and ensemble static ratings indicates that the FFE does not occur because the subjects are exposed to more information in the video clips, as the ensemble stimuli contained the same images contained in the video clips. There are still two important differences between the ensemble stimuli and the videos. First, the ensemble static images were not foveally presented (in virtue of being simultaneous), whereas the videos were. Second, the videos contained motion. The following experiment addresses these by presenting videos whose motion is manipulated at the fovea.

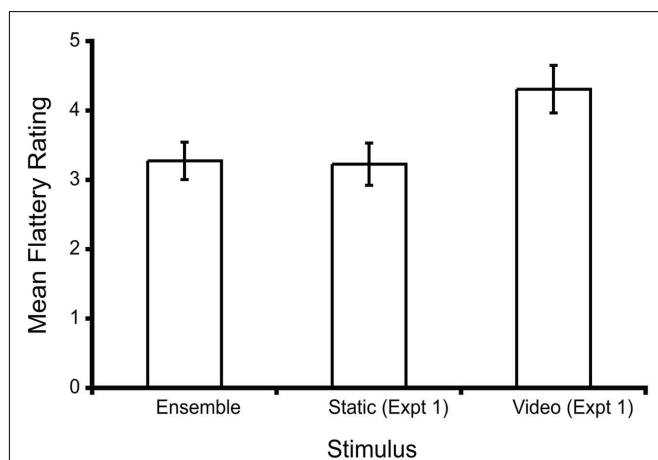


FIGURE 5 | Mean flattery ratings for the ensemble stimuli of Experiment 4. Also shown are the ratings for the component static images and corresponding video clips in Experiment 1. Error bars denote \pm SEM.

EXPERIMENT 5: IMAGE MOTION CHARACTERISTICS

The first four experiments suggest that motion may be a necessary condition for production of the FFE. Experiment 5 varied the image motion characteristics to explore the degree to which variations from natural facial motion may alter the FFE.

METHOD

Subjects

Five of the subjects from Experiment 1 participated in the study.

Stimuli

The stimuli were the 40 videos used in Experiment 1. Each video was presented under four conditions: forward motion, inverted forward motion, reversed motion, and scrambled frame sequence. The inverted forward motion conditions used the video clips from Experiment 2. The reversed motion videos were created by presenting the video clips in reverse. The scrambled frame sequence videos were created by presenting all frames contained within each video clip in a random sequence. All videos were standardized at 15 fps and 2 s duration.

Procedure

Subjects viewed the four video conditions in a random order and gave flattery ratings in the same manner described in the prior experiments.

RESULTS AND DISCUSSION

Mean flattery conditions were computed for each of the four stimulus conditions. These means are presented in **Figure 6**. It is apparent in the figure that the ratings for the scrambled condition were less than those for the other three conditions. In comparison, the other three conditions, which contained continuous motion, evoked more similar ratings. An ANOVA revealed a statistically significant difference among the four conditions depicted in **Figure 6** ($F_{3,16} = 16.82$, $p < 0.001$). *Post hoc* comparisons between all individual conditions indicated that the statistically significant differences were those between the scrambled condition and each of the other three conditions ($p < 0.001$ in each case).

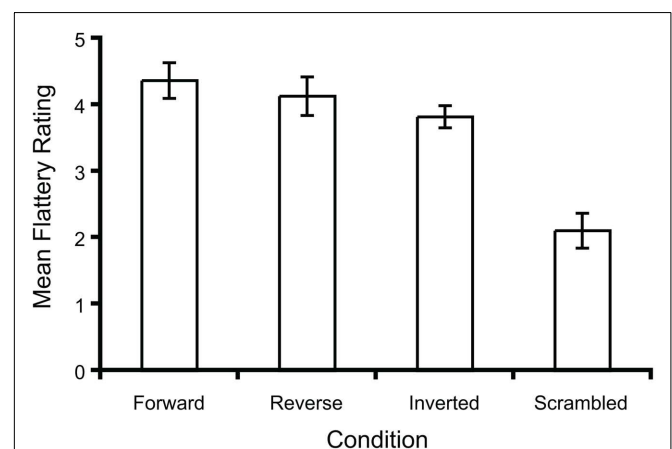


FIGURE 6 | Mean flattery ratings for the stimuli of Experiment 5. Error bars denote \pm SEM.

Because ratings for the video with scrambled frame sequence were significantly different than those for the other three conditions, it is evident that an important component of the FFE is that frames must be presented in sequence, whether it be forward motion, inverted forward motion, or reversed motion. The results suggest that the FFE requires the motion of faces in a continuous sequence.

EXPERIMENT 6: THE INVERTED FFE

The subjects in Experiments 2 through 5 had all participated in Experiment 1. To expand the generalizability of the findings, Experiment 6 was conducted as nearly a replication of Experiment 2, in which upright and inverted stimuli were examined using a within-subjects design. Only naïve subjects participated in this experiment. Additionally, the ratings in the dependent measure were switched from “flattery” to “attractiveness” to determine if similar results would be obtained. In this way, the FFE may be more directly compared to other results investigating attractiveness.

METHOD

Subjects

Ten undergraduate students (three males, seven females), aged 18–22 participated. All were naïve to the hypotheses of the research.

Stimuli

The stimuli were the 40 videos and 1200 static frames created for Experiment 1, and inverted copies of them.

Procedure

The procedure was similar to that of Experiment 2, with all videos and static frames presented in random, intermixed order, and subjects providing attractiveness ratings for the person in each video or static frame using a seven-point scale.

RESULTS AND DISCUSSION

A two (image: static vs. video) by two (orientation: upright vs. inverted) repeated-measures ANOVA was conducted for the attractiveness ratings. Both main effects of stimulus ($M_{\text{vid}} = 4.6$; $M_{\text{static}} = 3.9$; $F_{1,10} = 11.76$, $p < 0.01$) and orientation ($M_{\text{upright}} = 4.3$; $M_{\text{inverted}} = 4.15$; $F_{1,10} = 5.27$, $p < 0.05$) were statistically significant (see **Figure 7**). The interaction of the two factors is not significant ($F_{1,10} = 3.59$, $p = 0.09$). Thus, as in Experiment 2, the FFE was obtained despite inversion of the images, although the size of the effect is somewhat mitigated in the inverted condition.

The results obtained with the upright stimuli are similar to those of the other experiments, although with an entirely new set of observers. In the present experiment, however, the response measure was perceived attractiveness, rather than perception of how flattering they were, connecting it with past work examining attractiveness ratings (e.g., Knappmeyer et al., 2002).

GENERAL DISCUSSION

The current studies examined the FFE, wherein static images of people in mid-action are typically perceived as less flattering than videos from which the static images were derived. The FFE may

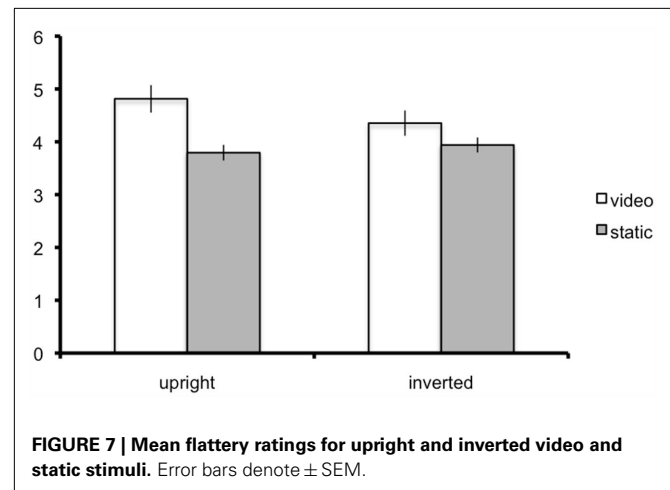


FIGURE 7 | Mean flattery ratings for upright and inverted video and static stimuli. Error bars denote \pm SEM.

be related to observations in prior research by Knappmeyer et al. (2002), which suggested that facial motion might influence ratings of attractiveness for computer-generated facial stimuli. While the report of Knappmeyer et al. (2002) may be consistent with the FFE, the FFE is distinct in that it demonstrates that moving faces are more flattering than most of the still images *composing* them.

The stimuli in our experiment were natural movies and therefore have ecological validity. One might worry, however, that the degree of familiarity with the actors in the videos might influence the FFE. To address this, we confirmed that the FFE occurred for the vast majority of the movie clips, for every observer (see Results and Discussion). Thus, the FFE occurs regardless of the observer's familiarity with the actor portrayed. One might also wonder whether the FFE is restricted to facial motion produced during speaking. Whether the FFE occurs for non-speaking videos of faces is an interesting question worthy of future investigation; however, because a large percentage of our exposure to faces in natural situations is to talking faces, the scope, and prevalence of the FFE is broad.

Overall, we are able to make the following conclusions: (1) the FFE exists in naturalistic settings; (2) the FFE can still occur with inversion; (3) the FFE is not due to superior memory for particularly distinct or attractive faces; (4) the FFE is not due to the fact that video clips contain more information than is contained in any singular, static image; and (5) the FFE requires the continuous motion of faces.

The FFE can be interpreted as poor temporal sampling of facial expression (at least during speech) – that is, the rapidly changing facial configurations during speech are temporally low-pass filtered by perception. This finding is consistent with prior work showing that observers preferentially represented the average emotion from a set of faces presented over time over any singular face within that set (Haberman et al., 2009), and also with the temporal integration of expression perception, which extends across several video frames in our displays (Arnold and Lipp, 2011). Furthermore, a number of studies show that average faces can appear more attractive (Langlois and Roggman, 1990; Rhodes and Tremewan, 1996; Winkielman et al., 2006), supporting the idea that the visual system is deriving summary information from the videos

over time. The FFE arguably is adaptive in that it minimizes the negative affective response to the brief facial variations during speech. It may also be useful in integration of facial perception across brief interruptions such as blinks.

This phenomenon may also explain why photography of people is so challenging. In general, photographers must capture the individual in an ideal position, under proper lighting conditions, and from the proper angle, and even then often apply post-processing (e.g., filters) to make the still image appear “right.”

Evolutionarily, humans have long been exposed to dynamic facial expressions and biological motion information, while static images such as those from cameras are relatively recent inventions. This could help explain the contribution of facial motion to enhanced recognition that has been reported in several studies (e.g., Pike et al., 1997; Hill and Johnston, 2001; Stone, 2001). Although mechanisms of face recognition are well characterized, what constitutes beauty, in general, or an “attractive” face, in particular, is less clear. One intriguing possibility is that faces are perceived as more attractive when they optimally drive the neural

mechanisms of face recognition (Winkielman et al., 2003, 2006). For example, the brain may more readily process dynamic faces than static faces. Although we *can* recognize a static image of a face, it may not be as easy to recognize as a face in natural motion, and this processing efficiency may be tantamount to enhanced attractiveness. This, of course, is speculation because what counts as attractive, as well as the neural correlate of attractiveness, remain elusive. Nevertheless, it is an interesting parallel, and it may explain why photography of faces is so difficult to master and why people anecdotally believe they look worse in photographs. Thus, the FFE offers one hypothesis as to how we decide whether something is deemed attractive, and also highlights the importance of using dynamic stimuli in addition to static images in the study of facial recognition.

ACKNOWLEDGMENTS

This research was supported in part by NSF grant 074689 to David Whitney. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

- Arnold, A. H., and Lipp, O. V. (2011). Discrepant integration times for upright and inverted faces. *Perception* 40, 989–999.
- Farah, M. J., Wilson, K. D., Drain, M., and Tanaka, J. N. (1997). What is “special” about face perception? *Psychol. Rev.* 105, 482–498.
- Haberman, J., Harp, T., and Whitney, D. (2009). Averaging facial expression over time. *J. Vis.* 9, 1–13.
- Haberman, J., and Whitney, D. (2007). Rapid extraction of mean and gender from sets of faces. *Curr. Biol.* 17, R751–R753.
- Haberman, J., and Whitney, D. (2009). Seeing the mean: ensemble coding for sets of faces. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 718–734.
- Hill, H., and Johnston, A. (2001). Categorizing sex and identity from the biological motion of faces. *Curr. Biol.* 11, 880–885.
- Knappmeyer, B., Thornton, I. M., Etcoff, N., and Bülthoff, H. H. (2002). Facial motion and the perception of facial attractiveness. *J. Vis.* 2, 612.
- Langlois, J. H., and Roggman, L. A. (1990). Attractive faces are only average. *Psychol. Sci.* 1, 115–121.
- Morrison, E. R., Gralowski, L., Campbell, N., and Penton-Voak, I. S. (2007). Facial movement varies by sex and is related to attractiveness. *Evol. Hum. Behav.* 27, 186–192.
- Pike, G. E., Kemp, R., Towell, T., and Phillips, K. (1997). Recognizing moving faces: the relative contribution of motion and perspective view information. *Vis. Cogn.* 4, 409–437.
- Rhodes, G., and Tremewan, T. (1996). Averageness, exaggeration, and facial attractiveness. *Psychol. Sci.* 7, 105–110.
- Rubenstein, A. J. (2005). Variation in perceived attractiveness. *Psychol. Sci.* 16, 759–762.
- Stone, J. (2001). Face recognition: when a nod is better than a wink. *Curr. Biol.* 11, 663–664.
- Thompson, P. (1980). Margaret Thatcher: a new illusion. *Perception* 9, 482–484.
- Winkielman, P., Halberstadt, J., Fazendeiro, T., and Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychol. Sci.* 17, 799–806.
- Winkielman, P., Schwarz, N., Fazendeiro, T., and Reber, R. (2003). “The hedonic marking of processing fluency: implications for evaluative judgment,” in *The Psychology of Evaluation: Affective Processes in Cognition and Emotion*, eds J. Musch and K. C. Klauer (Mahwah, NJ: Lawrence Erlbaum Associates), 189–217.
- Yin, R. K. (1969). Looking at upside-down faces. *J. Exp. Psychol.* 81, 141–145.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 September 2011; accepted: 18 January 2012; published online: 20 February 2012.

Citation: Post RB, Haberman J, Iwaki L and Whitney D (2012) The frozen face effect: why static photographs may not do you justice. *Front. Psychology* 3:22. doi: 10.3389/fpsyg.2012.00022

This article was submitted to *Frontiers in Cognition*, a specialty of *Frontiers in Psychology*.

Copyright © 2012 Post, Haberman, Iwaki and Whitney. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.